A Steady-State Kalman Filter for Assimilating Data from a Single Polar Orbiting Satellite*

DON BANFIELD AND ANDREW P. INGERSOLL

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California

CHRISTIAN L. KEPPENNE

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 20 July 1993, in final form 3 August 1994)

ABSTRACT

A steady-state scheme for data assimilation in the context of a single, short period (relative to a day), sun-synchronous, polar-orbiting satellite is examined. If the satellite takes observations continuously, the gains, which are the weights for blending observations and predictions together, are steady in time. For a linear system forced by random noise, the optimal steady-state gains (Wiener gains) are equivalent to those of a Kalman filter. Computing the Kalman gains increases the computational cost of the model by a large factor, but computing the Wiener gains does not. The latter are computed by iteration using prior estimates of the gains to assimilate simulated observations of one run of the model, termed "truth," into another run termed "prediction." At each stage, the prediction errors form the basis for the next estimate of the gains. Steady state is achieved after three or four iterations. Further simplification is achieved by making the gains depend on longitudinal distance from the observation point, not on absolute longitude. For a single-layer primitive equation model, the scheme works well even if only the mass field is observed but not the velocity field. Although the scheme was developed for *Mars Observer*, it should be applicable to data retrieved from Earth atmosphere satellites, for example, *UARS*.

1. Introduction

Kalman filtering is a recipe for blending predictions of a model with observations of a physical system to obtain an optimal estimate of the current system state (e.g., Gelb 1974). Unfortunately, Kalman filtering increases the computational burden of the model by a factor of order M, the number of state variables in the model system. This burden is prohibitive for atmospheric general circulation models (GCMs), where M is 10^5 or greater. In addition Kalman assimilation is often numerically unstable for large systems (Bierman 1977), so Kalman filtering has been an unreachable goal for atmospheric GCMs (Ghil et al. 1981; Miller 1986; Ghil 1989; Cohn and Parrish 1991; Daley 1992b).

The problem arises because the prediction error covariance matrix, whose dimension is $M \times M$, must be advanced forward in time as the system evolves. This time-dependence arises because the observing pattern, the observation error, and the model error can all be time dependent. But when these quantities (and thus the prediction error covariance matrix) are constant in time, the Kalman filter approaches a steady state. In such a case, the computational burden can be reduced to a factor of order 1; that is, the model runs almost as fast with data assimilation as without it because the prediction error covariance matrix need not be advanced in time. Moreover, the performance of such a steady-state Kalman filter is optimal, at least for linear constant coefficient systems forced by white noise. Steady-state Kalman filters (Wiener filters) have much in common with optimal interpolation (OI; see Bengtsson and Gustavsson 1971; Rutherford 1972; Bergman 1979), which is much used in operational forecasting. But currently operational observing systems have little in common with a system comprising only a single polar orbiting satellite. The present paper is an attempt to describe and evaluate an application to a single-satellite observing system.

Knowledge of the prediction error covariance matrix is necessary for computing the optimal weights (gains) that are used for blending the predictions and the observations together. Our scheme uses an estimate of the gains to assimilate data during a simulated flight of the spacecraft, from which an estimate of the prediction

* Contribution No. 5304 from the Division of Geological and Planetary Sciences, California Institute of Technology.

Corresponding author address: Don Banfield, Cornell University, 420 Space Sciences, Ithaca, NY 14853.

error covariance matrix is obtained. The simulation uses one run of a model as "truth" and another run for prediction. Simulated observations, with realistic errors, are made from the truth run. The resulting data are then assimilated into the prediction run using the current best estimate of the gains. The errors are measured by subtracting the truth field from the prediction field, and the error products are collected as sums. After integration for a suitable length of time, the sums are turned into averages and a new estimate of the prediction error covariance matrix is obtained. This leads to a new estimate of the gains, and the iteration is repeated. Convergence is complete in three or four iterations. The steady-state gains are then used with real data for as long as the observing pattern and the statistical properties of the system remain constant in time.

There are advantages and disadvantages to using a model as truth. An advantage is that the model's truth is known, whereas that of the real atmosphere is not. Knowing the truth allows us to bypass costly forward integration of the prediction error covariance matrix: instead of trying to compute the errors, we can measure them, which requires much less computer time. The disadvantage is that the model truth is not the real truth, which is the atmosphere itself. Since we are using a model both for prediction and for truth, our estimates of the error would tend to be low. To counter this identical twin problem, we add system noise to the truth run but not to the prediction run. The result provides an approximation to the true model error.

Our scheme was developed for use with *Mars Observer*. Although the spacecraft failed before going into orbit, the goal was to produce a continuous record of global weather over the planned two-year lifetime of the mission. The Mars GCM (MGCM) developed at Ames Research Center (Pollack et al. 1990; Haberle et al. 1992) was chosen as the model. Like all GCMs, it computes the state vector (temperature, pressure, wind, dust, and water vapor on a finite-difference global grid) at time $t + \Delta t$ from that at time t. Our plan was to assimilate data continuously, that is, to use observations taken during each integration step of the model to modify the state vector at that time step. The amount of modification takes into account the expected error of the observation and the expected error of the prediction. The output of the model after the assimilation step becomes the current best estimate of the global weather. The model then integrates this estimate forward to the next time step, during which new observations are collected and the process is repeated.

Mars Observer was designed for a sun-synchronous polar orbit with a period of 0.08 sols (Martian days). While Mars spins, the orbit stays fixed at the same local time of day, so successive dayside equator crossings move west by slightly less than 30 degrees. In one sol the spacecraft covers the globe with 12–13 equally spaced (in longitude) polar passes on both dayside and nightside. For our purposes, a steady-state observing system is one where the same instruments are taking data from the same point in the orbit on each polar pass. Our first objective was to assimilate data from one instrument, the Pressure Modulator Infrared Radiometer (PMIRR), which obtains profiles of temperature, water vapor, and dust to 80-km altitude every 115 km downtrack, that is, along the track of the spacecraft (McCleese et al. 1992). Since the spacecraft moves at 3.3 km s^{-1}, more than 10 profiles are obtained during each 7-minute time step of the MGCM.

The foundations of our assimilation technique are expanded upon in the next section, showing its roots in standard OI theory and its relation to Kalman and Wiener filters. We discuss the assumptions that we have made in the details of our assimilation technique. We also introduce a method of determining our gain functions, which we call Wiener gains. We have experimented with a number of simple (and not so simple) models to verify our assumptions. The remainder of the paper is devoted to detailing those models and the conclusions drawn from them. We first use a linear Rossby wave model that has one dimension (longitude) and one dynamical variable (horizontal streamfunction). We demonstrate the effectiveness of our technique of deriving the Wiener gains on this model. We then use a 2D one-variable model to show that the gain functions can be made functions of latitude as well as latitudinal and longitudinal distance between the spacecraft and the analysis point. With this model, we also investigate the severity of ignoring the serial correlations of observation errors. The next model is a three-variable Rossby wave/gravity wave model in one dimension. Only one variable, geopotential, is observed. The wind vector is analyzed from its correlation with geopotential. Furthermore, with this model we test whether topography is a serious perturbation to the Wiener gains, and, thus, whether we can ignore it. Finally, we implement our scheme on a nonlinear spherical shallow-water model, tuned to simulate Mars. The full implementation on the MGCM will be described in a later paper.

2. Assimilation gains

a. Background theory

The key to data assimilation lies in determining the weighting of the observations in computing an analysis. These weights, or gain functions, specify the changes to be applied to the prediction at an analysis point as a function of the difference between the prediction and observation at the observation point. This is mathematically written as

$$\Psi_k^a = \Psi_k^p + \sum_{j=1}^{N} \alpha_{kj}(\Psi_j^o - \Psi_j^p), \qquad (1)$$

where Ψ_k is an element of the state vector (at a point k) and the α_{kj} are the gain function coefficients speci-

fying how the discrepancy between the observation and prediction at point j should affect the state variables at point k. There are N observations in this update. The superscripts a, p, and o represent the analysis, prediction, and observation state vectors, respectively.

Optimal interpolation (OI) is based on minimizing the analysis errors that result from blending model predictions with observational data (Bengtsson and Gustavsson 1971; Rutherford 1972; Bergman 1979). Optimal interpolation tries to determine the values of α_{kj} that give the best estimate of the true state of the physical system. For a proper derivation of OI, the reader is referred to Rutherford (1972). The assumptions that one makes in the definition of OI are that the physical system and the observing system are statistically steady and that the observational errors are uncorrelated with the prediction errors. Minimizing the errors made in the analysis state vector yields a matrix equation for the gain functions as a function of the observation error covariance and the prediction error covariance,

$$\sum_{j=1}^{N} (\overline{\epsilon_i^p \epsilon_j^p} + \overline{\epsilon_i^o \epsilon_j^o})\alpha_{kj} = \overline{\epsilon_k^p \epsilon_i^p}, \qquad (2)$$

where the ϵ_i^p are prediction errors at point i and the ϵ_i^o are observation errors at point i. The overbars indicate averaging over many realizations. This equation is solved for the gain functions by inverting the matrix on the left side.

The above equations are expressed in a form that implicitly handles a large number of observations at once, assimilating them in parallel. For Earth, observations from the global synoptic network are typically assimilated at 0, 6, 12, and 18 h UTC (e.g., Baker et al. 1987; DiMego 1988). Satellite observations are usually simply moved to the synoptic times where they are assimilated along with data from the synoptic network. For Mars, however, and to some extent for the Earth's upper atmosphere, the data are all asynoptic. Different longitudes are observed at different times. Our approach is to assimilate data serially, at the time step nearest to when they were taken.

If the observation errors are uncorrelated, that is, the observation error covariance matrix $\overline{\epsilon_i^o \epsilon_j^o}$ is diagonal, then each observation can be handled separately. Equation (2), expressing the gain function for each observation, then becomes a simple division:

$$\alpha_{ki} = \frac{\overline{\epsilon_k^p \epsilon_i^p}}{\overline{\epsilon_i^p \epsilon_i^p} + \sigma_i^2}, \qquad (3)$$

where σ_i^2 is the error variance for observation i. While in general it is not true that observation errors are uncorrelated (Daley 1992a), we will show later that making this assumption presents little loss in performance for an application like *Mars Observer*. Jazwinski (1970, 198) discusses processing the components of the observation vector one by one when the observation error covariance is diagonal and there is no change in the state vector due to the dynamics. He states, "That this is equivalent to processing the whole observation vector at once is conceptually clear and can be proved algebraically (very tedious!)." Jazwinski does not give a proof, but we outline one in the Appendix.

Equation (1) relates an observation at point j to the kth element of the state vector, Ψ_k. Efficient but suboptimal forms of OI are obtained by assuming that the prediction error correlations are compact in space, symmetric about the observation point, and have Gaussian or other simple dependence on distance (e.g., Baker et al. 1987). In this paper k includes all elements, so that the assimilation step is global in extent. We make no assumptions about the spatial form of the prediction error covariance matrix. Further, we make no assumptions (e.g., geostrophic balance) about how different variables (e.g., wind and pressure) are related. Instead, we rely on the prediction error covariances to establish the relation.

The equivalence of these gains to the gains of a steady-state Kalman filter is easily seen by examining the equation for the discrete Kalman gain function (expressed in matrix form):

$$\mathbf{K} = \mathbf{P}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \qquad (4)$$

where $\mathbf{K}$ is the Kalman gain function matrix, $\mathbf{P}\mathbf{H}^{\mathrm{T}}$ is the prediction error covariance matrix times the transpose of the observation matrix, and $\mathbf{R}$ is the observational error covariance matrix (Gelb 1974). This equation is simply a more general form of (2), where the matrix $\mathbf{H}$ contains the transform between the observations and the state vector elements. Thus, our technique differs from Kalman only in the way in which the prediction error covariance matrix is determined. Kalman filtering advances it mathematically in time with the state vector, while in our technique, it is measured via models and stored for later use with real data.

b. *Gain simplifications*

In order that the gains be constant in time, the atmosphere and the observation system must be statistically steady. This means that we must have different gains for different seasons, and perhaps also for periods when the atmosphere is clear and periods when it is dusty. In addition, the characteristics of the instrument must not change; that is, the observation error must be constant in time. And finally, the observing pattern must be steady.

The last condition needs some explanation. We need separate gains for each latitude of observation, because the coverage of the poles is different from that of the equator. For example, points near the poles are observed every orbit, whereas points near the equator are observed once per day. But if the instrument is taking data all the time, then each time the spacecraft passes a certain latitude the positions of past observations rel-

ative to the spacecraft is the same. In other words, the observing pattern is constant. This satisfies the criterion for a steady-state Kalman filter. Even if the instrument is off during part of the orbit, the gains will be steady provided the on–off cycle is a function of latitude only.

In principle, the gains should also be functions of the longitude of observation. Relative to a scale height, the topography of Mars is significantly greater than the Earth's (Esposito et al. 1992), so the weather at one longitude is systematically different from that at another. Thus, the prediction error covariance matrix should depend not only on the longitudinal distance from the observation point j to the analysis point k, but also on the absolute longitude of both. We have chosen to ignore this latter dependence. Below, we show that even for Mars, with its great topographic relief, the assumption presents an insignificant loss of performance.

c. *Gain function recipe*

As shown in (2) and (3), the gains α_{kj} depend on the error covariances. The observation error depends on instrument noise and on retrieval uncertainty—that associated with converting calibrated radiances into vertical profiles of physical quantities. The retrieval process introduces correlations within each profile, so the observational error should be a matrix whose dimension is the number of points in the profile. Such considerations are beyond the scope of this paper, however, since we are not discussing the implementation into the 3D MGCM. In most of our simpler models presented herein, we assume that the observational errors are uncorrelated from one horizontal location to the next. The observational error variance is an input parameter to the assimilation scheme.

The other quantity that appears in (2) and (3) is the prediction error covariance, which we obtain by iteration. The first estimate comes from the weather covariances, the spatial correlations of the weather fluctuations, which are the differences between the instantaneous state vector and its time mean. These covariances are much larger than the prediction error covariances obtained with the optimal gains, but they have a similar spatial structure.

The iteration process has been described earlier. With an estimate of the error covariances, one determines the gains using (2) or (3). One then assimilates some synthetic observations of the truth run into a prediction run, keeping track of the prediction errors. From these a new prediction error covariance is computed, and the process repeats. We do not mathematically show the convergence of this method, but have tested it on a number of different models, and all have converged in a few iterations. Furthermore, we have compared it to Kalman filters implemented on linear models, and the gain functions converge to the same functions as the Kalman gains in steady state.

3. 1D one-variable model

In this section, we introduce the iterative procedure yielding the Wiener gains with an application to a 1D Rossby wave model. We used a 1D Rossby wave model on a latitude circle written in a geopotential-like variable, with damping and random forcing. The equation of motion is

$$\left(\frac{\partial}{\partial t}+\frac{1}{\tau}\right)\left(\frac{\partial^2}{\partial x^2}-\frac{1}{L_D^2}\right)\Psi+\beta\frac{\partial\Psi}{\partial x}=F, \quad (5)$$

where F is the random forcing, taken to be white noise in space and time, and β, L_D, and τ were chosen to represent possible Mars-like conditions at the equator. Here x is longitude with period 2π, and t is time with one sol being 2π; L_D is the radius of deformation. We choose $L_D = 1/3$, about one-third of Mars' radius in the model's dimensionless units. It emphasizes wavenumber three Rossby waves, like those reported in Barnes (1980, 1981) from the *Viking Lander* data; τ was taken to be 8π, corresponding to a decay time of 4 sols (Barnes 1980, 1981; Banfield et al. 1994), and β is 2 in these dimensionless units. The equation yields a spectrum of Rossby waves. Wavenumbers 1, 2, and 3 have the largest amplitudes and a coherence time of order 4 sols. We decomposed the system into 12 spectral modes of Fourier sine and cosine functions plus a constant term, with 25 terms in the representation. Observations of the state were degraded with Gaussian noise, which had an amplitude of 5% of the average amplitude of the system. We modeled our observing system after *Mars Observer,* which would have crossed the equator on a descending node about 12 times per sol. Thus, the observing pattern was simply an observation about every $1/12$ of a sol. This was also the model time step as we analytically advanced the models in time from one observation to the next. The simplicity of the system allowed us to *analytically* integrate the equation of motion from time t to time $t + \Delta t$, at which point the forcing changes discontinuously and a new observation is made. There is no leapfrog time step; the system remains first order in time and is computationally stable. We are thus able to derive explicit formulations for the transition and noise-forcing matrices of the Kalman formalism.

We implemented a Kalman filter for this system, which is depicted as it begins to assimilate data in Fig. 1. The vertical bars in the truth model column are the actual observations made under the spacecraft, which is drifting westward, making one cycle in a sol. They simply represent the value of the observed geopotential under the spacecraft, as measured from the dotted line which is the x axis, with some artificial observational noise. The Kalman filter very quickly adjusts the estimate to the truth run in roughly a sol. Similarly, the shape of the Kalman gains can be seen to change over that time, as the prediction error covariances adjust to the new influx of information. The gain functions

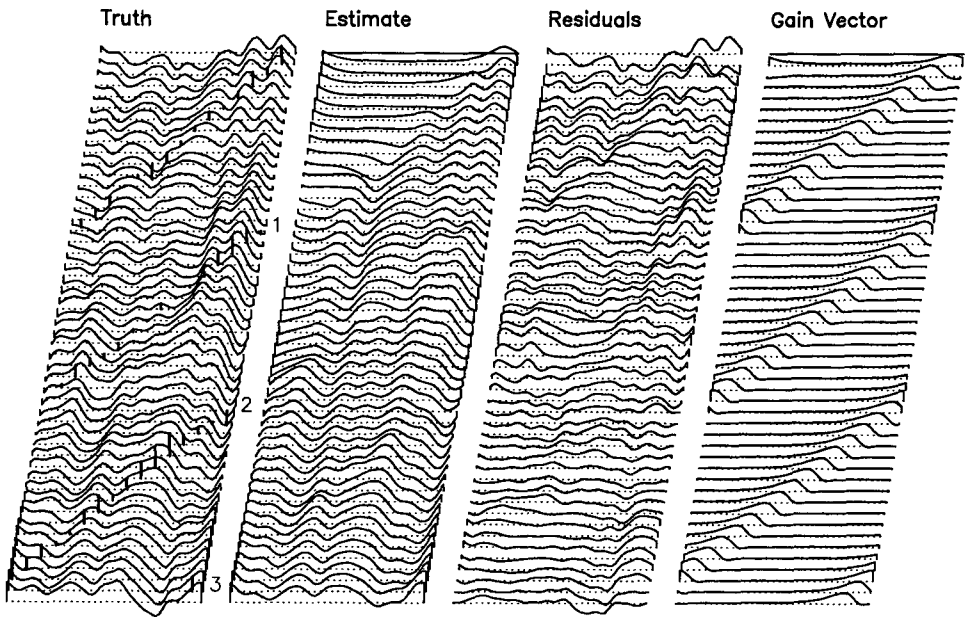

FIG. 1. One-dimensional one-variable model's evolution over about 3 sols, with time advancing downward and each column representing one quantity over the full domain of the model. The leftmost column shows the state of the truth run, or what is taken to be the real atmospheric state. Observations, represented by the vertical bars from the dotted zero line to the curve, are being taken by a satellite that progresses westward with time. The second column corresponds to an estimate of the true state made by a Kalman filter. The third column shows the difference between the truth and the estimate. Finally, the fourth column shows the Kalman gains used to weight the observation taken at that time step. Note the speed with which the Kalman filter improves the estimate, and how the shape of the gain function changes over the first sol.

quickly settle to a steady state that is a function of the longitudinal difference between the spacecraft and the analysis point only. In other words, the gain functions have a constant shape in the reference frame of the spacecraft.

Initially the Kalman gains are symmetric about the observation point and fairly broad. However, over the course of the next sol, they exhibit a shorter correlation length and more asymmetry from east to west. The asymmetry arises because the location just to the east of the spacecraft was visited one orbit (0.08 sols) earlier, while the location to the west of the spacecraft was visited 12 orbits (one sol) earlier. Additionally, β causes information to be advected westward, adding to the asymmetry. The steady-state Kalman gain follows the spacecraft as it travels westward. It is depicted again as the solid line in Fig. 2. Asymmetry in the shape of the gain is particularly evident in this figure, where the spacecraft is taken to be at $x = 0$.

The analyses are all done in the spectral coordinates of the model, and thus the gain functions relate an observation in space to a correction of a spectral mode. The gains in Fig. 2 are obtained from the spectral coefficient gains by a linear transformation. This method is used throughout the paper, including the more complex models that follow. It is a natural choice for these models as they are formulated in spectral space, and it is currently in use in operational schemes (Parrish and Derber 1992).

The other two curves in Fig. 2 are the zeroth and second iteration of the Wiener gains for this system. The zeroth iteration is computed from (3) using the total weather covariance as the prediction error covariance. This is necessarily a suboptimal gain function, as it is the equivalent of the prediction errors made using no predictive model at all. It does, however, roughly correspond to the western half of the Kalman gain function, perhaps indicating that that side of the Kalman gain is controlled by the correlation length scale of this model's weather. After some 1000 sols of model time, the statistics were clean enough on the prediction error covariances in this iteration to yield the new (first iterate) gain functions. This procedure was then repeated to generate the second iteration gain functions, the dotted line in Fig. 2. The latter are strikingly similar to the Kalman gains, and perform just as well as they do. Further iterations change the gain functions insignificantly from either the second iterate, or the Kalman gain. Therefore, we find that this iterative technique does in fact converge to the steady-state Kalman gains, or the true Wiener gains, for this simple model.

Figure 3 shows the predictive performance during the first 10 sols of assimilation using the Kalman gains,

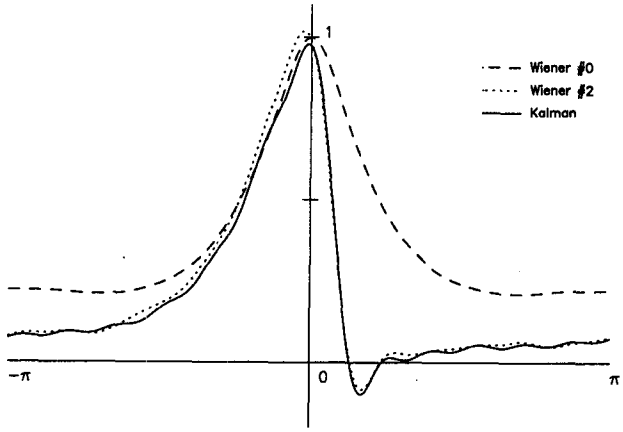

FIG. 2. Three gain functions for the 1D one-variable model, the zeroth iterate Wiener gain, the second iterate Wiener gain, and the steady state Kalman gain. The spacecraft is at $x = 0$. A gain of 1.0 means that the observation is solely used to make the analysis, while a gain of 0.0 means that only the prediction is used in the analysis. The similarity between the second iteration Wiener gains and the Kalman gains shows the near optimal nature of the iterated Wiener gains.

the zeroth iterate Wiener gains, and the second iterate Wiener gains. Random weather variations make these curves much less smooth for any particular 10 sols, but this plot is the result of averaging 1000 independent 10 sol start-up periods for the different filters. The prediction models were initialized to a zero state for all of these runs, so the initial errors were of the same order of magnitude as the model weather variance. The errors are normalized to the field's average variance, a normalized error of 1.0 corresponding to essentially no accuracy in the analysis, and a normalized error of 0.0 meaning that the predictive model exactly tracks the truth model's trajectory. Again, the solid curve corresponds to using the Kalman gains, which were able to bring the error variance down to about 25% of the model's weather variance in just over one sol. The dashed line corresponds to using the zeroth iterate Wiener gains, which took two or three sols to reach a steady state with an error variance of about 46%. The dotted line corresponds to using the second iterate Wiener gains, the performance of which is almost indistinguishable from the Kalman gains.

This is a good model for comparing the computational burdens of the Kalman filter versus the constant Wiener filter. Every sol of assimilation using the Kalman filter on this model took three times as long as the Wiener filter to execute. The Kalman filter has the extra overhead of propagating the prediction error covariance matrix through every time step. For more complex models, this step becomes prohibitively slow because it scales as M^2, the square of the number of elements in the state vector. In our models, the Wiener filter's computational requirements scale as M, the same as in OI. Therefore, we expect our technique to present about the same computational burden as OI, and significantly less than Kalman filtering.

4. 2D one-variable model

To test the validity of this technique in a model with varying weather statistics from north to south, we used a 2D one-variable Rossby wave model. This system is described by

$$\left(\frac{\partial}{\partial t}+\frac{1}{\tau}\right)\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}-\frac{1}{L_D^2}\right)\Psi+\beta\frac{\partial\Psi}{\partial x}=F, \quad (6)$$

where the only differences from the 1D model are that the state vector now consists of the components of a 2D Fourier transform. We used somewhat lower resolution for this model to reduce the computational burden; zonal wavenumbers of zero to seven, and meridional wavenumbers of one to five satisfying a no-normal-flow boundary condition. We implemented a Kalman filter for this model and also derived the iterated Wiener filters. The spacecraft orbit was modeled as traveling once per sol around the planet to the west, and making about 12 north–south passes in that time. The gain functions were kept in the reference frame of the spacecraft as separate functions for each latitude that the spacecraft observed, because the weather and prediction error covariances are assumed to depend on latitude only. Also, the gain functions were kept as *global* functions, relating the observation point to analysis points everywhere in the domain. This is the same

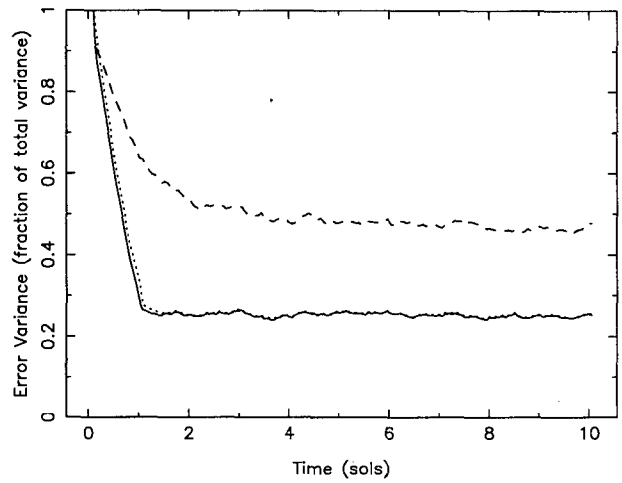

FIG. 3. Performance of the three gain functions of Fig. 2. The ordinate is the variance of the analysis errors from the model's true state. The error variances are normalized by the model's weather variance, so that an error variance of 1.0 is correctly analyzing essentially none of the model's weather variance. The curves run for 10 sols, and were all started from a zero state at the initial time. The dashed curve is for the zeroth iterate Wiener gains, which performs relatively poorly. The other two curves correspond to the iterated Wiener gains and the Kalman gains, which are nearly identical in performance and clearly superior to the zeroth iterate Wiener gains.

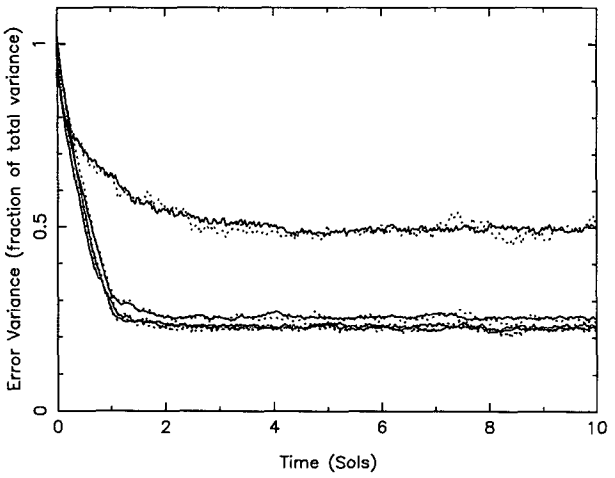

FIG. 4. Performance of the zeroth, first, second, and third gain functions for the 2D one-variable model both with and without serially correlated observation errors. This is for observation errors of 5% of the total weather variance of the model. The solid lines represent the runs without serially correlated observation errors, and the dotted lines represent the corresponding runs with serially correlated observation errors. The error variance is normalized in the same way as in Fig. 3. The zeroth iterate yielded the solid and dotted lines asymptoting to about 0.5, while the subsequent iterates yielded the lower error variances. Note the similarity between each solid line and its corresponding dotted line.

as in the 1D model, but with a 2D domain and separate gain functions for different spacecraft latitudes. This implementation of the iterated Wiener filter worked just as well as the 1D implementation, with similar improvements in performance between the zeroth iterate gain function and the later iterates. Furthermore, the performance difference between the Kalman gains and the later iterations on the Wiener gains were indistinguishable. Finally, the shape of the Wiener gains was almost indistinguishable from the Kalman gains after only two iterations.

This model is also a good one with which to examine the impact of serially correlated observation errors. In this series of experiments, the observations were degraded with noise of the form $n_i = \eta n_{i-1} + (1 - \eta^2)^{1/2} r_i$, where n_i is the noise of the ith observation; η is a number less than 1, and r_i is a random number. The correlation coefficient between the noise at position i and that at position $i - m$ is η^m. The n_i have the same mean and variance as the r_i. We process these observations as before, ignoring the fact that the errors are correlated, and see what effect an increase in η has on the performance of the assimilation scheme. The serial correlations were modeled to have roughly the same meridional structure as the weather fluctuations themselves. For 10 observations in each north–south pass, this means a serial correlation of about 0.75 from one observation to the next. We modeled the northernmost observation as uncorrelated with observations preceding it.

Presented in Fig. 4 is the performance of the zeroth, first, second, and third iteration gain functions with an observation error of 5% relative to the weather fluctuations. The solid lines represent the performance without serially correlated observation errors $\eta = 0$, and the dotted lines represent those with serially correlated observation errors $\eta = 0.75$. The performance loss in ignoring these serial correlations is apparently negligible for this application, which in structure resembles the scenario of *Mars Observer*. The observation errors comprise such a small fraction of the prediction errors for this problem that whether they are serially correlated or not is insignificant. However, if the observation errors were of greater amplitude, this is less true.

The PMIRR instrument of *Mars Observer* would have yielded atmospheric radiances that could have been inverted for temperatures with an accuracy of about 1–2 K (McCleese et al. 1992). The atmospheric waves on Mars represent variations on the order of 5 K (e.g., Barnes 1980). Therefore, a more reasonable estimate of the observation errors for the *Mars Observer* scenario is about 40%. Figure 5 shows the decrease in performance when serially correlated observation errors are present for the case of 40% errors. It shows a more notable decrease in performance than the 5% case, indicating that observation errors contribute more strongly to the prediction errors in this case. Nevertheless, the performance loss is small compared to the performance difference between the zeroth iterate gain function and subsequent iterates of the gains functions. Because of this, we believe that it is an adequate assumption to ignore the possibility of serial correlations in the observation errors for the *Mars Observer*

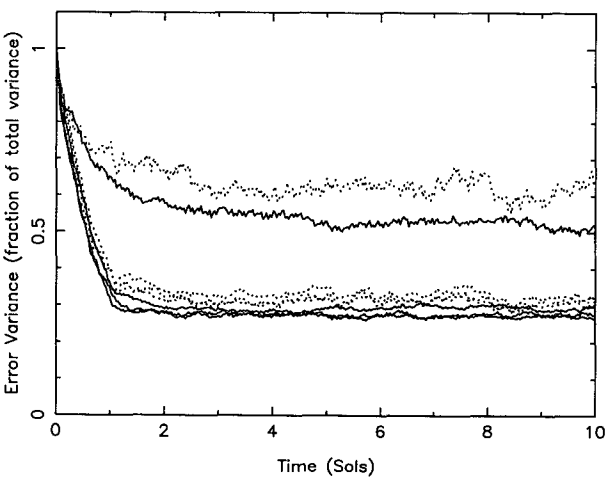

FIG. 5. Same as Fig. 4 but for observation errors of 40% of the total weather variance. This value is similar to that which had been expected for *Mars Observer*. Note that the later iterates still perform better than the zeroth iterate, even when there are serially correlated observation errors. This suggests that ignoring serial correlations of the observation errors does not represent a serious loss in performance.

scenario. However, if the observation errors are still greater, the assumption clearly will break down. Figure 6 shows the performance difference for a case with 200% observation errors. In this case, ignoring the serial correlations causes performance losses that are much greater than the performance increases found between the zeroth and subsequent iterations.

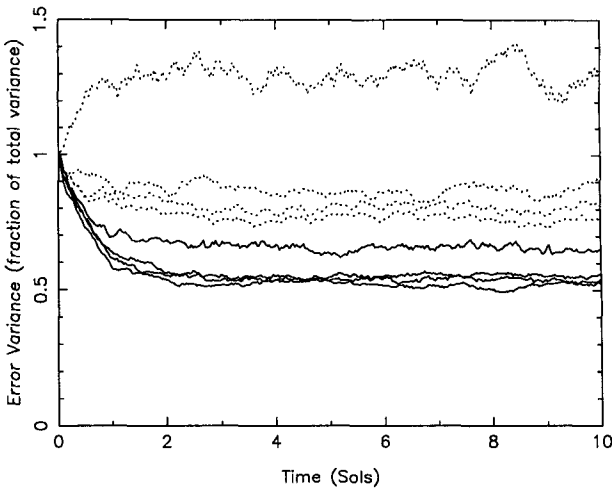

FIG. 6. Same as Fig. 4 but for observation errors of 200% of the total weather variance. In this extreme case, the runs with serially correlated observation errors perform significantly worse than those without. The zeroth iterate's performance is actually worse than no assimilation at all. In such a case, serial correlations of observation errors cannot be ignored.

5. 1D three-variable model

To test the performance of the Wiener gains on the primitive equation system, we applied them to the following linearized 1D shallow-water model. The objective was to observe only the mass field and see how well the mass and velocity fields follow the truth run:

$$\left(\frac{\partial}{\partial t}+\frac{1}{\tau}\right)\Psi + f\chi = F$$

$$\left(\frac{\partial}{\partial t}+\frac{1}{\tau}\right)\chi - f\Psi + \Phi = G$$

$$\left(\frac{\partial}{\partial t}+\frac{1}{\tau}\right)\Phi + c^2\frac{\partial^2\chi}{\partial x^2} = H. \qquad (7)$$

In (7) Ψ is the streamfunction, χ is the velocity potential, and Φ is the geopotential; τ is again a damping time (~4 sols), f is the Coriolis parameter (taken to be at 30°N), and beta is neglected in order to simplify the algebra for Kalman filter implementation. Here c^2 is the critical wave speed, with a value of 1/9. With these settings, the radius of deformation, $L_D = c/f$, is again 1/3. The forcing functions, F, G, and H are red in their time and space spectra (Lorentzians centered at zero frequency and zero wavenumber, with half-power points at a period of 4 sols and a longitudinal wavenumber of 3), so as not to force the gravity waves too strongly. The relative amplitudes of the forcings are chosen to mimic the nonlinear terms, which have been dropped from the system. Thus, $F \sim H$ and $G \sim F/4$, because the amplitude of χ is one-quarter that of Ψ and Φ when the forcing period is 4 sols. This model is like the simple 1D one-variable model with the addition that it also admits fast modes. The time evolution of the model's three variables during an 8-sol run of this model is shown in Fig. 7. The geopotential and streamfunction fields are nearly in geostrophic balance. They also closely resemble the geopotential field of the simple one-variable model discussed above. However, the velocity potential field is dominated by rapidly propagating east- and westbound gravity waves. It is smaller in amplitude than the other two fields by a factor of 5, reflecting the fact that gravity waves are not strongly forced in this model and that geostrophic balance dominates. The dynamical result of neglecting β is that the slow modes do not propagate, although they drift randomly under the influence of the forcing. However, Rossby waves propagate rather slowly even when β is included, so we feel that this is not a crucial difference. Geopotential only is observed, and a 5% error has been added to the observations as before.

Again, we tested the iterative Wiener filter technique on this model and compared the results to a real Kalman filter. The results were similar to those obtained with the two Rossby wave models. After a few iterations, the Wiener gains were almost indistinguishable from the Kalman gains, and performed equally well. Figure 8 shows the Kalman gains, and the zeroth and fifth iteration Wiener gains. In this case there are three gain functions, because they relate an observation in geopotential to changes in all three dynamical variables. One can still notice the asymmetry between the east and west sides of the subspacecraft point first noticed with the 1D one variable model. In this case, the asymmetry is not as dramatic, probably because this model has no β, which likely serves to enhance east–west asymmetries.

The performance increase from the symmetric zeroth iterate gain function to the fifth iterate is not as drastic as with the simpler models, but still worth noting. Figures 9a, 9b, and 9c depict the magnitude of the prediction error variance for the first 10 sols of assimilation for the streamfunction, velocity potential, and geopotential fields, respectively. The predictive runs were initialized to a zero state in all cases. The zeroth iterate (dashed line) reaches a steady state in about 3 sols, predicting all but 18% of the variance of the model in streamfunction, 16% in geopotential, and essentially nothing of the velocity potential field (>98% unexplained variance). The fifth iterate Wiener gains (dotted line) reach a steady state in only about 2 sols, pre-

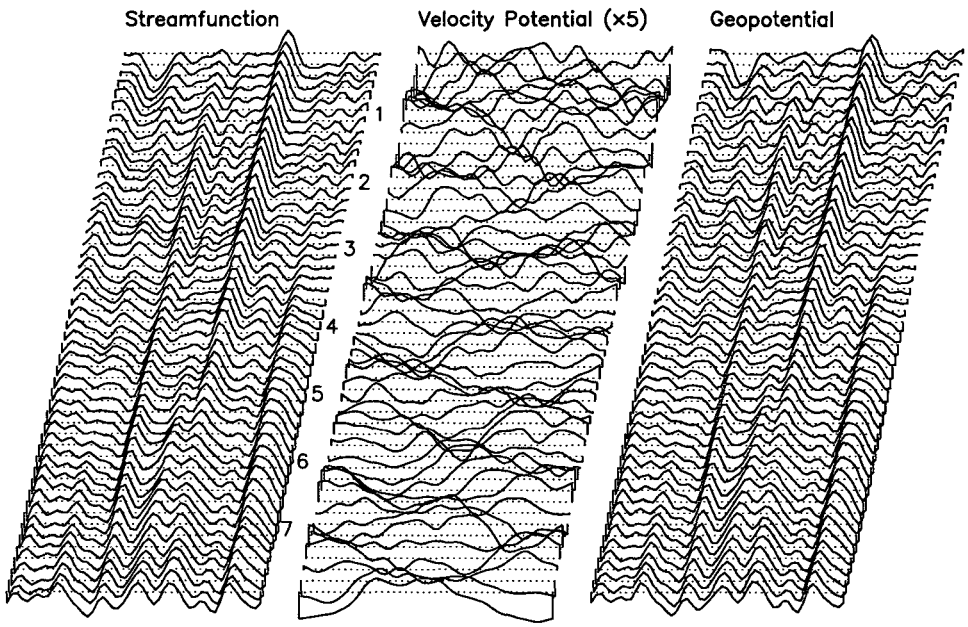

FIG. 7. Evolution of the 1D three-variable model over an 8-sol interval. Again, time advances downward, and each column shows a dynamical quantity over the whole domain of the model. The streamfunction and geopotential evolve slowly and exhibit a great degree of geostrophy, while the velocity potential is dominated by quickly propagating gravity waves.

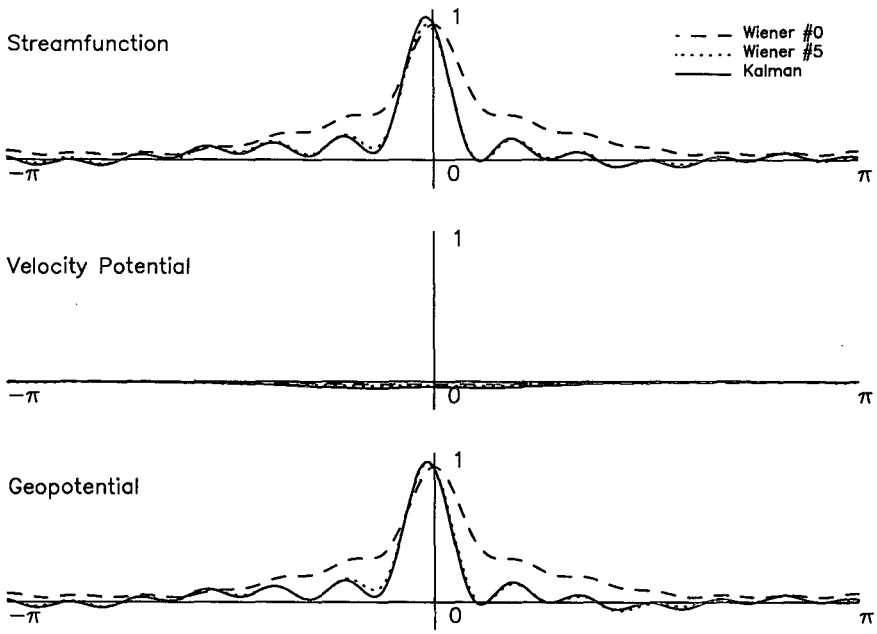

FIG. 8. Gain functions for the 1D three-variable model, each relating an observation of geopotential to an analysis weight in one of the dynamical variables. As in the one-variable model, the iterated Wiener gains are nearly indistinguishable from the Kalman gains and longitudinally asymmetric. The gain functions for streamfunction and geopotential are similar due to the high degree of geostrophy present in the model. The low values of the velocity potential gain functions are indicative of the independent nature of geopotential and velocity potential in this model.

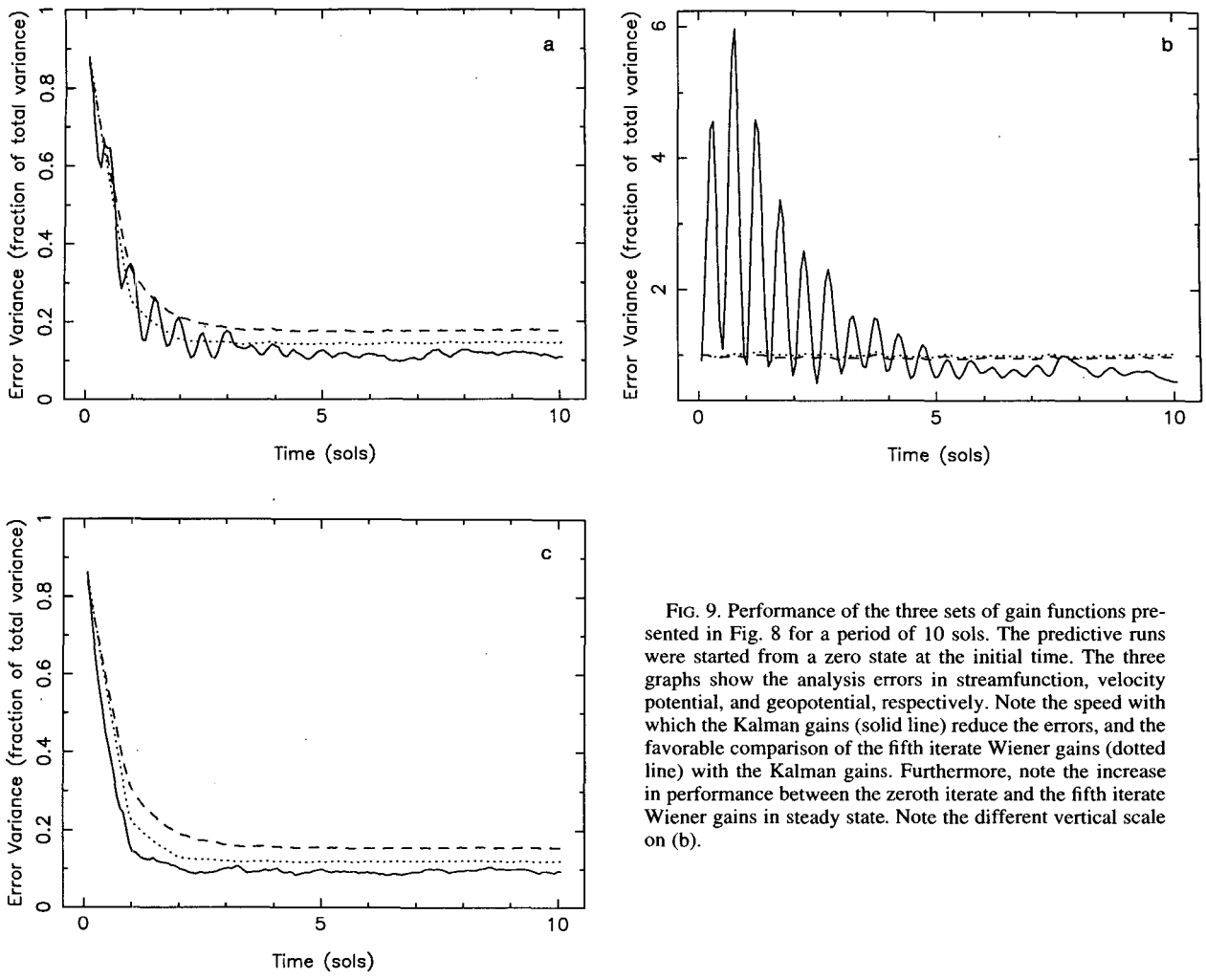

FIG. 9. Performance of the three sets of gain functions presented in Fig. 8 for a period of 10 sols. The predictive runs were started from a zero state at the initial time. The three graphs show the analysis errors in streamfunction, velocity potential, and geopotential, respectively. Note the speed with which the Kalman gains (solid line) reduce the errors, and the favorable comparison of the fifth iterate Wiener gains (dotted line) with the Kalman gains. Furthermore, note the increase in performance between the zeroth iterate and the fifth iterate Wiener gains in steady state. Note the different vertical scale on (b).

dicting all but about 15% of streamfunction, 12% of geopotential, and again essentially none of velocity potential (>98% unexplained variance). The Kalman filter (solid line) produces some strong oscillations at first, especially in the velocity potential field, probably as a result of improperly specified initial prediction error covariances. However, at least in the geopotential field, it settles down in about 2 sols. When the filter reaches its steady state, it predicts all but about 9% of geopotential variance, about 11% of streamfunction variance, and appears to predict a small percent of the velocity potential variance (~70% unexplained variance).

That none of the sets of gain functions is able to effectively predict the velocity potential is not surprising because the only observed field is geopotential, which is dominated by Rossby waves, while the velocity potential is dominated by gravity waves. The lack of communication between these fields is also reflected in the much smaller amplitudes for the velocity potential gain function. The fact that streamfunction and geopotential gain functions are similar in shape and predictive ability simply reflects a high degree of geostrophy in the model.

This model is interesting in that it demonstrates how one can effectively use observations of one dynamical variable to predict the others. Cross-correlations between an observed variable and other dynamical variables are commonly used in operational assimilation schemes. It is these cross-correlations that allow one to use information about one observed dynamical variable to guide the trajectory of the entire dynamical system through phase space. *Mars Observer* would have returned temperature soundings of the Martian atmosphere, and would have no access to direct wind measurements. It is then important to be able to predict the full dynamical state of the system from measurements of only one variable. This scheme, like currently operational OI schemes (e.g., Parrish and Derber 1992), makes no assumptions about dynamical balance in the cross correlations relating different variables. Rather, we use the prediction error covariances relating the sys-

tem variables with each other. This should give the right relation between the dynamical variables of the model system, without the difficulties geostrophy presents near the equator.

We neglected β in this model in order to be able to easily solve the equations of motion analytically, and thus easily cast the system into a Kalman filter formulation. However, in anticipation of real applications of this technique with finite-difference models (GCMs), we implemented this model in a finite-difference formulation as well. We used a leapfrog time-stepping scheme to advance the dynamics, thereby changing the system from first order to second order in time. The new state vector (at time $t + \Delta t$) is computed from the current and past state vectors (times t and $t - \Delta t$). We applied the same corrections to the past state vector as to the current one. Applying the correction to only the current state vector excited the computational mode in the integration.

Transforming this model into a finite-difference model also allowed us to add β and topography to the model. We have been ignoring topography in assuming that the gain functions are not functions of the absolute longitude of the spacecraft. Adding topography to the model is a way of checking this assumption. With these changes, the equations of the system become

$$\left(\frac{\partial}{\partial t} + \frac{1}{\tau}\right)\Psi + f\chi = F$$

$$\left(\frac{\partial}{\partial t} + \frac{1}{\tau}\right)\chi - f\Psi + \Phi = G$$

$$\left(\frac{\partial}{\partial t} + \frac{1}{\tau}\right)\Phi + \frac{\partial}{\partial x}\left(c^2(x)\frac{\partial \chi}{\partial x}\right) - \beta L^2 \frac{\partial \Phi}{\partial x} = H, \quad (8)$$

where $c^2(x)$ represents the topography through $c^2(x) = gh(x)$, gravity times the local depth of the layer. These expressions are not directly derivable from the shallow water equations, but they have the same behavior at both low and high frequencies. They have the further property of energy conservation in the absence of forcing and damping. We used a Galerkin scheme with a finite number of Fourier modes as the set of basis functions. This scheme conserves energy when the leapfrog time step is sufficiently small. We chose the function graphed in Fig. 10 as the layer depth. It is roughly 13 km minus the meridionally averaged topography of Mars, scaled to yield a high degree of modulation on $c^2(x)$. Since the topographic relief of Mars is quite dramatic (on the order of an atmospheric scale height), this should be a good test of the relevance of specifying the gain functions in the spacecraft's frame of reference.

We include the effect of topography on the dynamics both for the truth runs, which have random forcing, and for the prediction runs, which are forced only by assimilated observations. However, we ignore the effect of topography on the gain functions when we assume that they are the same for all spacecraft longitudes. With this assumption, we derived the longitudinally averaged iterated Wiener gains in the spacecraft's frame of reference and evaluated the predictive performance of these gain functions. Then, we included the effect of topography by having separate gain functions for the different spacecraft longitudes, and again derived a set of Wiener gains. These latter gain functions explicitly included information about the local topography, and thus should have performed better than the former gain functions.

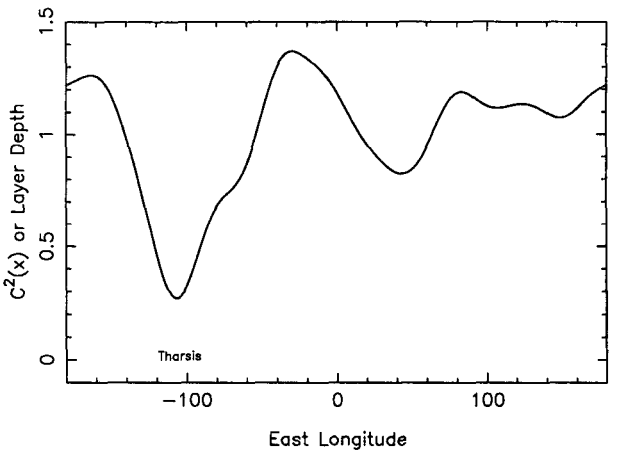

FIG. 10. Layer depth of the 1D three-variable model as a function of longitude. It corresponds to the actual Martian topography, subtracted from a reference of 13 km, meridionally averaged and scaled to yield a high degree of modulation on the quantity c^2. The Tharsis plateau of Mars shows up quite clearly at around −110°. It appears as a low value because this is layer *depth*, as opposed to surface height.

The solid line in Fig. 11 represents the gain functions derived ignoring the presence of topography, assuming that the same function can be used for all spacecraft longitudes. The dashed line shows a gain function derived with topography explicitly included, one of a set which is different for each spacecraft longitude. This particular gain function is for the longitude with the greatest topographic relief, the Mars Tharsis region. As one might expect, this is the location where the two sets of gain functions differ the most. However, even in this region, the differences are not dramatic. Throughout most of the rest of the domain, the differences are considerably smaller. The impact of these differences in the gain functions is most evident in Fig. 12, which shows the errors made in the geopotential field during the first 10 sols of assimilation using different gain functions. The dotted line corresponds to using the zeroth iterate Wiener gains. It takes 4 sols to reach a steady-state in which about 30% of the geopotential variance is unexplained. The solid line corresponds to the third iterate Wiener gains derived ignoring topography. The dashed line is the result of us-

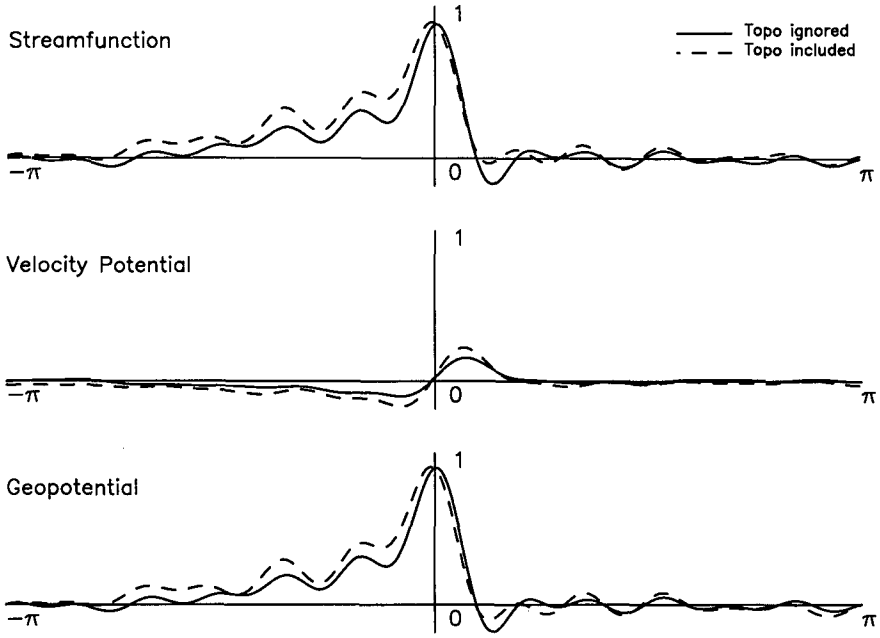

FIG. 11. Effect of topography on the gain functions. The dashed line is for the Tharsis region, where the topographic relief is greatest. The gain function for this region differs the most from the gain function derived by ignoring topography (solid line). The differences are not astounding in this example, and are much smaller at other longitudes.

ing the gain functions including topography. Their performance is essentially indistinguishable from the gain functions derived without including topography.

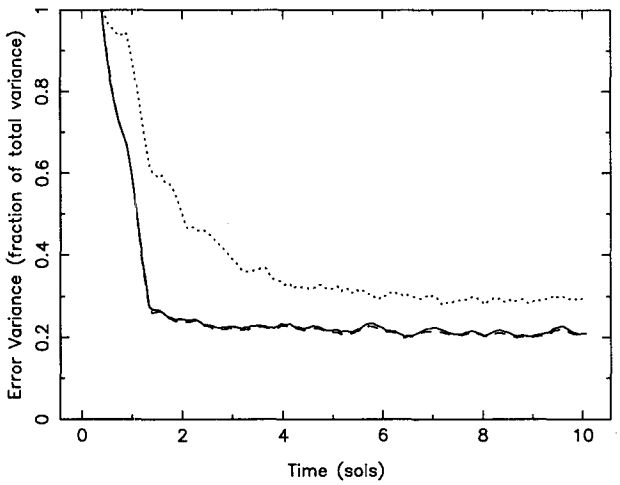

FIG. 12. Performance difference in analyzing the geopotential field using gain functions derived ignoring the presence of topography and gain functions derived explicitly including topographic effects, for a 10 sol run. The predictive model was started from a zero state. The solid line corresponds to the gain functions derived ignoring topography, while the dashed line corresponds to the gain functions with topography included. The performance difference is insignificant, particularly when compared to the performance of the zeroth iterate Wiener gains (dotted line).

Both reach steady state in one to two sols, and can explain all but about 22% of the geopotential variance. Similar results were found for the other two state variables, with streamfunction being equivalently well predicted as geopotential (~22% unexplained variance) and velocity potential remaining essentially unpredicted (~88% unexplained variance).

Thus, while local differences can be found between the gain functions that accounted for topography and those that did not, the performance difference is insignificant. The gain functions, which included topography, outperformed those without topography by less than 1% of explained variance, while the standard symmetric zeroth iterate gain functions performed roughly 8% worse than either of the Wiener gains. Therefore, we believe that it is acceptable to ignore topography and any other satellite longitudinal dependence of the gain functions.

6. Spherical shallow-water model

Since our first operational application of the Wiener filter will be on a full-scale multilayer Martian GCM (Pollack et al. 1981, 1990; Haberle et al. 1992), a high-resolution, single-layer primitive equation model appeared appropriate to serve as the last milestone in our testing hierarchy. We have, thus, derived the Wiener gains for the single-layer spherical shallow-water model discussed in Keppenne (1992). For this more complex model, we used slightly different values for

two parameters than those used in the simpler models. At a latitude of 30° the radius of deformation, L_D, is roughly the same as the planetary radius, ~3 times greater than in our earlier examples. The relaxation time is 10 sols, which is a factor of 2.5 times greater than in our earlier examples. A realistic mountain field is taken from the topography of the Pollack et al. (1990) Martian GCM, whose horizontal resolution is equivalent to the one used here. This model's global behavior is fairly representative of the Martian atmosphere, and locally, the statistics of its geopotential time series match well those of the surface pressure data at the sites of *Viking Landers I* and *II*.

The modeled spacecraft trajectory was chosen to obey a polar sun-synchronous orbit, with a sampling rate of 1000 geopotential observations per sol—about one observation every 250 km of spacecraft downtrack motion. Since *Mars Observer* would have observed the vertical structure of the temperature field over the entire depth of the Martian atmosphere, one single geopotential observation here amounts to a snapshot of the vertical temperature profile in the case of a multilayer GCM. As in the other experiments, the observations were degraded with 5% Gaussian noise, and the model was randomly forced with red noise so that the truth and predictive runs were not identical twins. The variance of the random forcing term that affects the time evolution of each spectral coefficient was set equal to that of the corresponding nonlinear term.

Since this is a nonlinear model, there is no analytic solution to the optimal filtering problem. Although an implementation of the Kalman filter by linearizing the model about its continuously evolving state is conceivable, such an implementation would overwhelm even the fastest present-generation supercomputers. We thus restricted our data assimilation applications on this model to the derivation of the corresponding steady-state Wiener gains. Three iterations were sufficient for the Wiener gains to reach a quasi-steady-state, that is, a fourth iteration with the predictive model assimilating data using the third iterate gain functions resulted in virtually no change from the third to fourth iterate. Each set of gain functions was computed from the statistics of the second half of a 2000-sol prediction run, continuously forced with geopotential observations of a comparatively long history tape of a truth run.

To successfully initialize the iterative procedure, it was necessary to scale down the zeroth iterate gain functions. When unscaled, these gain functions induced the predictive run's trajectory to oscillate wildly about the truth and in some cases "blow up." This undesirable behavior arose probably because the weather statistics, which were used in the zeroth iterate Wiener gains, differ more substantially from the steady-state error statistics than in the experiments with the linear models. The result could be an overestimation of the error covariances by the zeroth iterate Wiener gains. It is also possible that the zeroth iterate Wiener gains did not have the proper balance between variables (geopotential and wind, for example), and that the improper balance excited some of the model's unstable modes. Whatever the case, multiplying the zeroth iterate Wiener gains by a number from 0.1 to 0.5 was sufficient to alleviate this problem. The problem did not occur on subsequent iterations, and no further reductions of the gain functions were necessary. Once properly initialized, the iteration converged as before.

Contour maps of the longitudinally independent third iterate Wiener gain functions are shown in Fig. 13 in the spacecraft's frame of reference, with the subspacecraft grid point located along the center meridian at the appropriate latitude. These maps correspond to the Gaussian latitudes closest to 67.5°N, 22.5°N, 22.5°S, and 67.5°S. Presumably (since we could not check this assumption with an analytically derived Kalman filter), they show how geopotential observations should be weighted to optimally update the streamfunction (Figs. 13a–d), velocity potential (Figs. 13e–h), and geopotential field (Figs. 13i–l) to steer the predictive model toward the truth model.

Several points are worth noting about the shapes of these gain functions. First, the longitudinal asymmetry reported in our discussions of the linear models' gain functions is less noticeable, although it is still present (Figs. 13d and 13l). We are not sure of the cause of this reduction but suspect that it is due to the greater relaxation time used in this model (10 sols versus 4 sols previously). A longer relaxation time would result in less errors being built up over 1 sol. Thus, differences between the east and west sides of the observation point would also be reduced, and the gains would be more symmetric than in the previous models. Second, there is a substantial departure from geostrophic adjustment, especially in the Tropics (see Fig. 13b,c,j,k) where the Coriolis approximation cannot be made. The streamfunction weights are characterized by a dipole pattern at low latitudes (Figs. 13b and 13c). The geopotential weights (Figs. 13i–l) resemble more the Gaussian bell patterns of operational OI schemes. The velocity potential gain functions (Figs. 13e–h), are of much lesser magnitude than their streamfunction and geopotential counterparts because the velocity potential and geopotential fields are essentially uncorrelated. They display a more complicated organized structure that reflects long-distance correlations in the velocity potential error field.

The analysis performance of the third iterate Wiener gains are compared in Fig. 14, which shows the normalized streamfunction, velocity potential, and geopotential error variances for a 20-sol run of the predictive model against a history tape of the truth model. The predictive model was initialized with a zero initial state. After 20 sols, the third iterate Wiener gains give unexplained variances of 4%, 16%, and 3% in the streamfunction, velocity potential, and geopotential fields, respectively. Considerably larger residual errors

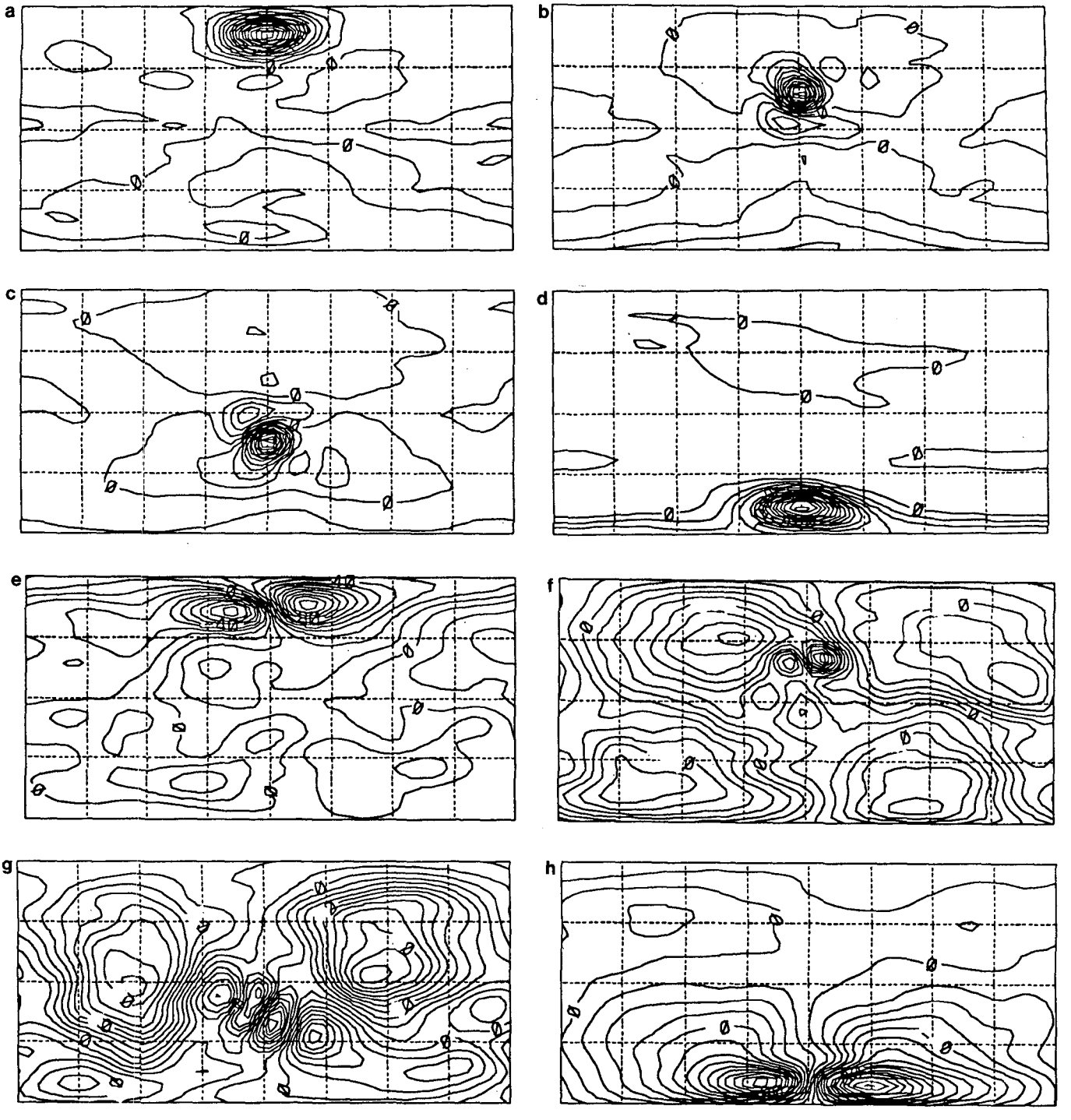

FIG. 13. Contour maps showing the longitudinally independent third iterate Wiener gains for the global nonlinear shallow water model in the spacecraft's frame of reference, with the subspacecraft grid point located along the center meridian at the appropriate latitude. These maps correspond to the Gaussian latitudes closest to 67.5°N, 22.5°N, 22.5°S, and 67.5°S. Figures 13a–d, e–h, and i–l correspond to the gain functions used to update the streamfunction, velocity potential, and geopotential fields, respectively, when observations are taken at these latitudes. These maps show how one should expect the theoretical Kalman gains to be spatially distributed, for this model of Mars' atmosphere when the observational pattern matches that of a single, sun-synchronous polar orbiting satellite.

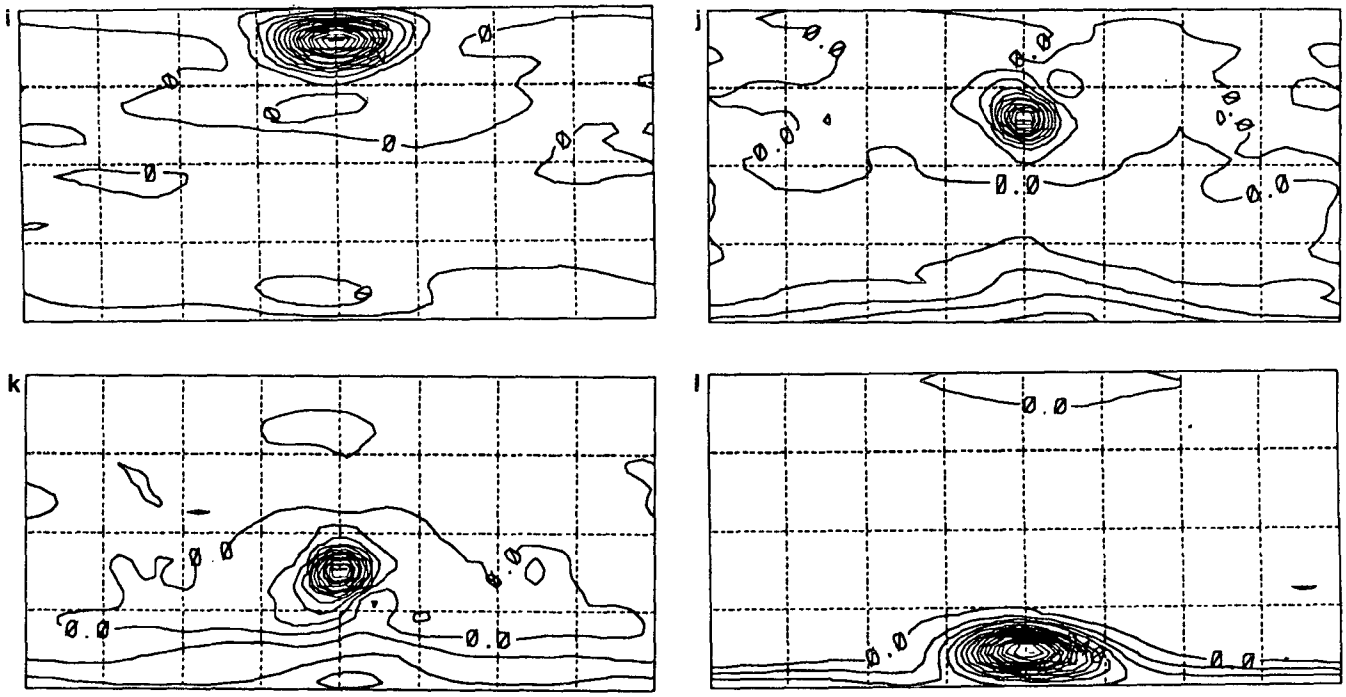

FIG. 13. (*Continued*)

remain when the zeroth iterate weights are used. Note that the streamfunction and geopotential errors decrease exponentially during the first day of the experiment and linearly thereafter, while the velocity potential errors decrease linearly from the start. The magnitude of the errors in assimilating with this model are smaller than those found with the other models tested. This is due to a difference in the strengths of the random forcings as compared to the simpler models. The random forcing completely dominated the behavior of the previous models, while in this model, the zonal flow and topography have a more significant effect on the climatology. Thus, the prediction runs and truth runs for this experiment were more similar than for the previous models, and analysis performance increased. The main result of the experiments with this model is simply that the assimilation scheme works on this advanced model.

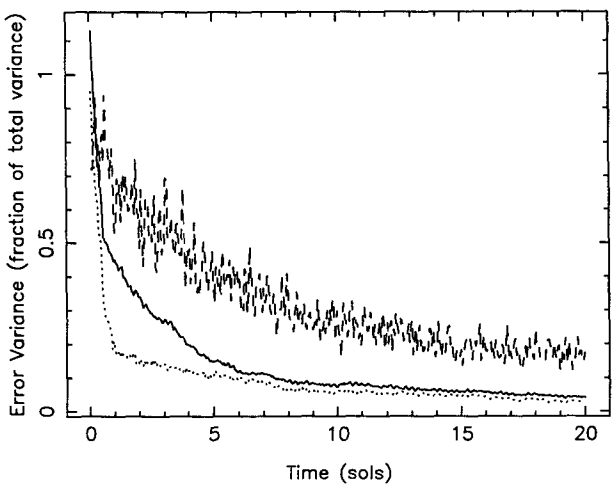

FIG. 14. Plot showing the performance of the third iterate Wiener gains derived with the global nonlinear shallow water model. The solid line shows the evolution of the normalized streamfunction error variances during a 20-sol prediction run against a history tape of a truth run, perturbed by random forcing terms. The dashed and dotted lines show the corresponding evolutions of the velocity potential and geopotential error variances, respectively.

7. Conclusions

The main result of this work is the development of a new approach to data assimilation to be used with a single polar-orbiting satellite like *Mars Observer.* It shares some of the advantages of OI and some of the advantages of Kalman filtering. As in OI the gains are constant in time, but they are equivalent to steady-state optimal Kalman gains. Because the gain functions are computed off-line, this technique has similar computational requirements as OI, and, thus, considerably less than Kalman filtering. We showed that the effects of longitudinal differences on the shape of the gain functions were not significant. This allows the optimal gain functions to be simplified to be a function of only the observational point's latitude. Furthermore, we showed that for an application like *Mars Observer,* serial cor-

relations in the observation errors could be safely ignored without a significant loss in predictive performance.

We also demonstrated the feasibility of using these gain functions on a suite of models of increasing complexity, approaching that of a full GCM. These models were observed in only one dynamical variable, as *Mars Observer* would have done for the Martian atmosphere, yet we were able to make predictions about all three dynamical variables. Furthermore, we verified the success of an iterative technique in determining these optimal Wiener gains with each model. The next step will be to employ these techniques on a full GCM-scale model, for the assimilation of data from the next Mars atmospheric mission, or perhaps of data from a single polar-orbiting stratospheric-sounding earth satellite.

APPENDIX

Equivalence of Serial and Simultaneous Assimilation

The proof is by induction: one shows that the equivalence holds for n observations provided it holds for $n-1$ observations. We define $\alpha_{kj}^{(n-1)}$ as the gain matrix, linking the observation at point j to the state vector at point k, in processing $n-1$ observations. Thus, for simultaneous assimilation of $n-1$ observations,

$$\Psi_k^{(n-1)} = \Psi_k^{(0)} + \sum_{j=1}^{n-1} \alpha_{kj}^{(n-1)}(\Psi_j^o - \Psi_j^{(0)}), \quad \text{(A1)}$$

where $\Psi_k^{(0)}$ is the state vector before the processing begins, and $\Psi_k^{(n-1)}$ is the state vector after $n-1$ observations have been processed. The $\alpha_{kj}^{(n-1)}$ are assumed to satisfy

$$\sum_{j=1}^{n-1} (r_{ij} + v_j\delta_{ij})\alpha_{kj}^{(n-1)} = r_{ik}. \quad \text{(A2)}$$

Here $r_{ij} = \overline{\epsilon_i^p \epsilon_j^p}$, where ϵ_i^p and ϵ_j^p are the prediction errors associated with $\Psi_i^{(0)}$ and $\Psi_j^{(0)}$, respectively. Also, $v_j = \sigma_j^2$ is the error variance of the jth observation, and δ_{ij} is the Kronecker delta. Equations (A1) and (A2) follow from the assumption that serial and simultaneous processing are equivalent for the first $n-1$ observations.

Serial processing can be written

$$\Psi_k^{(n)} = \Psi_k^{(n-1)} + \alpha_{kn}^{(n)}(\Psi_n^o - \Psi_n^{(n-1)}). \quad \text{(A3)}$$

Note that n denotes both the location of the observation and its order in the processing sequence. The $\alpha_{kn}^{(n)}$ are computed as in (3), but the errors ϵ_i^p and ϵ_j^p are the errors in $\Psi_i^{(n-1)}$ and $\Psi_j^{(n-1)}$, respectively. From (A1), these errors satisfy

$$\epsilon_k^{(n-1)} = \epsilon_k^{(0)} + \sum_{j=1}^{n-1} \alpha_{kj}^{(n-1)}(\epsilon_j^o - \epsilon_j^{(0)}). \quad \text{(A4)}$$

We substitute (A4) into (3), recognizing that the observation errors are uncorrelated with the prediction errors and that $r_{ij} = r_{ji}$. This yields

$$\alpha_{kn}^{(n)} = \frac{r_{kn} - \sum \alpha_{kj}^{(n-1)} r_{nj} - \sum \alpha_{nj}^{(n-1)} r_{jk} + \sum\sum \alpha_{ni}^{(n-1)} \alpha_{kj}^{(n-1)} (r_{ij} + v_j\delta_{ij})}{r_{nn} + v_n - 2\sum \alpha_{nj}^{(n-1)} r_{jn} + \sum\sum \alpha_{ni}^{(n-1)} \alpha_{nj}^{(n-1)} (r_{ij} + v_j\delta_{ij})}. \quad \text{(A5)}$$

The single sums are over j from 1 to $n-1$. The double sums are over i and j, both of which vary from 1 to $n-1$. With (A2) this becomes

$$\alpha_{kn}^{(n)} = \frac{r_{kn} - \sum \alpha_{kj}^{(n-1)} r_{nj}}{r_{nn} + v_n - \sum \alpha_{nj}^{(n-1)} r_{nj}}. \quad \text{(A6)}$$

Substituting (A1) into (A3), one can relate $\alpha_{kj}^{(n)}$ to $\alpha_{kj}^{(n-1)}$ for $j < n$:

$$\alpha_{kj}^{(n)} = \alpha_{kj}^{(n-1)} - \alpha_{kn}^{(n)} \alpha_{nj}^{(n-1)}. \quad \text{(A7)}$$

Equations (A6) and (A7) follow from (A3), the nth step of the serial assimilation. The remainder of the proof is showing that (A6) and (A7) are consistent with the expression for simultaneous assimilation, the n-dimensional version of (A2), that is, with $(n-1) \rightarrow n$. First separate out the nth row and nth column:

$$(r_{nn} + v_n)\alpha_{kn}^{(n)} + \sum_{j=1}^{n-1} r_{nj}\alpha_{kj}^{(n)} = r_{nk}, \quad \text{(A8)}$$

$$r_{in}\alpha_{kn}^{(n)} + \sum_{j=1}^{n-1} (r_{ij} + v_j\delta_{ij})\alpha_{kj}^{(n)} = r_{ik}, \quad \text{(A9)}$$

where i varies from 1 to $n-1$. Then substitute for $\alpha_{kj}^{(n)}$ in (A8) and (A9), using (A7). Equation (A9) becomes an identity and (A8) becomes (A6), which shows that the equations are consistent. In other words, serial and simultaneous assimilation are equivalent for n observations provided they are consistent for $n-1$ observations. Since they are equivalent for $n = 1$, they are equivalent for all n. A corollary is that the result of serial assimilation is independent of the order with which the observations are processed.

REFERENCES

Baker, W. E., S. C. Bloom, J. S. Woollen, M. S. Nestler, E. Brin, T. W. Schlatter, and G. W. Branstator, 1987: Experiments with a three-dimensional statistical objective analysis scheme using FGGE data. *Mon. Wea. Rev.*, **115,** 272–296.

Banfield, D., A. D. Toigo, A. P. Ingersoll, and D. A. Paige, 1994: Martian weather correlation length scales from Viking IRTM T15, *Icarus,* submitted.

Barnes, J. R., 1980: Time spectral analysis of midlatitude disturbances in the Martian atmosphere. *J. Atmos. Sci.*, **37,** 2002–2015.

——, 1981: Midlatitude disturbances in the Martian atmosphere: A second Mars year. *J. Atmos. Sci.*, **38,** 225–234.

Bengtsson, L., and N. Gustavsson, 1971: An experiment in the assimilation of data in dynamical analysis. *Tellus,* **23,** 328–336.

Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.*, **107,** 1423–1444.

Bierman, G. J., 1977: *Factorization Methods for Discrete Sequential Estimation.* Academic Press, 244 pp.

Bourke, W., 1972: An efficient, one-level, primitive-equation spectral model. *Mon. Wea. Rev.*, **100,** 683–689.

Cohn, S. E., and D. F. Parrish, 1991: The behavior of forecast error covariances for a Kalman filter in two dimensions. *Mon. Wea. Rev.*, **119,** 1757–1785.

Daley, R., 1992a: The effect of serially correlated observation and model error on atmospheric data assimilation. *Mon. Wea. Rev.*, **120,** 164–177.

——, 1992b: Estimating model-error covariances for application to atmospheric data assimilation. *Mon. Wea. Rev.*, **120,** 1735–1746.

DiMego, G. J., 1988: The National Meteorological Center regional analysis system. *Mon. Wea. Rev.*, **116,** 977–1000.

Esposito, P. B., W. B. Banerdt, G. F. Lindal, W. L. Sjogren, M. A. Slade, B. G. Bills, D. E. Smith, and G. Balmino, 1992: Gravity and topography. *Mars,* H. H. Kieffer, B. M. Jakosky, C. W. Snyder, and M. S. Matthews, Eds., The University of Arizona Press, 209–248.

Gelb, A., Ed., 1974: *Applied Optimal Estimation.* The MIT Press, 142–143.

Ghil, M., 1989: Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework. *Dyn. Atmos. Oceans,* **13,** 171–218.

——, S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. L. Bengtsson, M. Ghil, and E. Källén, Eds., *Dynamic Meteorology: Data Assimilation Methods.* Springer-Verlag, 139–224.

Haberle, R. M., J. B. Pollack, J. R. Barnes, R. W. Zurek, C. B. Leovy, J. R. Murphy, H. Lee, and J. Schaeffer, 1992: Mars atmospheric dynamics as simulated by the NASA Ames general-circulation model. 1. The zonal-mean circulation. *J. Geophys. Res.*, **98,** 3093–3123.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory.* Academic Press, 376 pp.

Keppenne, C. L., 1992: Orographically-forced oscillations in a dynamical model of the Martian atmosphere. *Icarus,* **100,** 598–607.

McCleese, D. J., R. D. Haskins, J. T. Schofield, R. W. Zurek, C. B. Leovy, D. A. Paige, and F. W. Taylor, 1992: Atmosphere and climate studies of Mars using the Mars Observer pressure modulator infrared radiometer. *J. Geophys. Res.*, **97,** 7735–7757.

Miller, R. N., 1986: Toward the application of the Kalman filter to regional open ocean modeling. *J. Phys. Oceanogr.*, **16,** 72–86.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120,** 1747–1763.

Pollack, J., C. Leovy, P. Greiman, and Y. Mintz, 1981: A Martian general circulation experiment with large topography. *J. Atmos. Sci.*, **38,** 3–29.

——, R. Haberle, J. Schaeffer, and H. Lee, 1990: Simulation of the general circulation of the Martian atmosphere. Part I: Polar processes. *J. Geophys. Res.*, **95,** 1447–1474.

Rutherford, I. D., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, **29,** 809–815.